The Effect of Naïve Theories of Intelligence on Metacognitive Monitoring Accuracy

Terry Purton

A report submitted as a partial requirement for the degree of Bachelor of Psychology with Honours at the University of Tasmania, 2016.

Statement of Sources

I declare that this report is my own original work, and that the contributions of others have been duly acknowledged.

Terry Purton

13 October, 2016

Acknowledgements

First and foremost, I would like to thank my supervisor, Dr. Matthew Palmer, for encouraging me to be autonomous in my learning, and for providing the knowledge, enthusiasm, and support that allowed me to do so. I am privileged to have had the opportunity to learn from such an experienced and passionate academic.

I would like to thank Dr. Christine Padgett, Dr. Cynthia Honan and Mr. Peter Tranent for allowing me into a number of their classes during participant recruitment. As well as Frances Parkes for her assistance with programming, and Nicole McCallum at Flinders University for assisting with data collection.

I would also like to give a special mention to Caitlin Gleeson, for sharing the responsibility of data collection, and for her emotional support. To my other fellow honours students, your company and support has made 2016 a year of fond memories. I wish you all the very best as we take this next step, and look forward to hearing about your future endeavours.

To my dear family, thank-you for believing in me and supporting me in all that I do. To Danni, thank-you for your great listening ear, and for never doubting me, even when I doubt myself. Lastly, to my friends' back home, thank-you for your messages of encouragement over this past year; they meant more than you will ever know.

Table of Contents

List of Tables and Figures

The Effect of Naïve Theories of Intelligence on Metacognitive Monitoring Accuracy

Terry Purton

Word count: 9,391

Abstract

Research suggests that individuals who believe their intelligence has the capacity to incrementally grow may make less accurate judgements of learning (JOLs) than those who believe intelligence is a fixed entity. JOLs represent a prediction of future recall, and have implications for real world study behaviour. The present study aimed to determine whether delaying JOLs could improve JOL accuracy for incremental theorists. It was hypothesised that there would be an interaction effect of theory of intelligence (TOI) and time on metacognitive accuracy, and that overall metacognitive accuracy would be better for delayed than immediate JOLs. 56 participants (46 females, 10 males) completed a paired-associate learning task, provided immediate and delayed JOLs and completed a cued-recall test and Dweck's (1999) TOI measure. There was no effect of TOI on either immediate or delayed JOLs (all p's > 0.1), negating the first hypothesis. The second hypothesis was supported, $F(1, 54) = 124.02$, $p < .001$, $d = 1.25$, replicating the delayed-JOL effect. These findings suggest that poor metacognition is not a concern for incremental theorists, and speak to the utility of learning techniques that use retrieval cues as a basis for study behaviour.

It has been established that individuals have implicit, naïve 'theories' of memory and cognition, which act as an interpretive framework for their metacognitive experiences during learning (Schwarz, 2004). Of relevance to the present research are the naïve theories that individuals hold about the malleability of their own intelligence (Romero, Master, Paunesku, Dweck, & Gross, 2014). Dweck and Leggett (1988) have identified two broad views about intelligence that people subscribe to; either that it is a fixed entity, or that it has the capacity to grow. The results of numerous studies suggest that these theories of intelligence (TOI) have differential effects on achievement-related behaviour, with an incremental (growth) view consistently associated with superior outcomes (see Claro, Paunesku, & Dweck 2016; Dweck, 1999).

However, recent evidence suggests that endorsing such a view may be problematic for the accuracy of metacognitive judgements (Miele, Finn, & Molden, 2011; Miele & Molden, 2010). Metacognition refers to the capacity of the human mind to appraise conscious thoughts and experiences, and use the information gained via this process to evaluate and adjust behaviour (Nelson, 1996). This practice of 'thinking about one's own thinking' has important implications for how people learn (Nelson, 1996). In particular, it allows an individual to monitor their understanding, and use this feedback to evaluate the quality of their memory for information, and predict the likelihood of future recall (Dunlosky, Serra, Matvey, & Rawson, 2005; Miele & Molden, 2010). These predictions are referred to as judgements of learning (JOLs), and are used to inform decisions about what, when, and how to study, playing a crucial role in academic achievement (Metcalfe & Finn, 2008).

The present study aims to investigate the assertion that having an incremental view of intelligence can lead to poor JOL accuracy. This is a relatively recent proposition, which, given the importance of accurate metacognitive judgements for optimal learning, makes it worthy of further investigation. Not only in order to determine the merits of the claim, but also to establish whether there exists a need for the development of learning strategies that account for differences in naive TOIs. It is hoped that the outcome of this research will contribute to a growing understanding of the ways in which psychological mechanisms influence learning (Dweck, 2015). Research which explores this concept is vital for the design of both effective study techniques and school curricula that account for individual differences in learning. It is thus of high practical importance, for policymakers, educators, and the wellbeing and success of all learners (Rattan, Savani, Chugh, & Dweck, 2015).

Theories of Intelligence, Goals and Effort Attributions

The Social Cognitive Model of Motivation (Dweck, 1999; Dweck & Leggett, 1988) suggests that there are two general sets of beliefs, or theories, that people have regarding the plasticity of their intelligence. *Entity theorists* believe that their intelligence is fixed at a given level and focus on proving this innate ability, whilst *incremental theorists* believe that their intelligence is malleable and focus on increasing it through hard work and effort (Dweck & Leggett, 1988). As such, Dweck & Leggett (1988) propose that entity theorists are motivated to demonstrate their ability through the pursuit of performance-based goals. These individuals focus on validating their intrinsic capabilities, and therefore choose to engage in tasks or activities which they believe allow them to demonstrate competence with minimal effort (Dweck, 1999). Conversely, incremental theorists are motivated to pursue learning goals, as they are focused on incrementally increasing their intelligence, and believe that engaging in novel, effortful, challenging tasks is the best way to achieve this (Dweck, 1999).

These conflicting conceptions about the meaning of effort have implications for how entity and incremental theorists respond to situations involving challenge and failure (Dweck, 1999). For incremental theorists, these situations signal a need to increase effort; to study harder, listen better, practice more (Dweck, 1999). This is typically referred to as a mastery-oriented response pattern, as the focus is on mastering the task at hand through increased output of physical and mental resources (Dweck & Leggett, 1988). This is adaptive in academic settings, as it encourages acceptance of the reality that difficulty is a part of the learning process, and therefore promotes persistence in the face of setbacks (Dweck & Leggett, 1988).

In contrast, for those with entity beliefs, the need for high effort indicates low ability on a task (Dweck and Leggett, 1988). Performance goals can encourage avoidance of learning opportunities, as they present a risk of failure and are thus not a means to achieve the central goal of demonstrating competence (Dweck & Leggett, 1988). As a result, entity theorists tend to use a helpless response pattern, often avoiding novel or challenging tasks, for fear that they will perform poorly, as well as being hesitant to persevere following failure (Dweck, 1999). It is evident then, that a person's naïve TOI has a profound impact on the types of tasks they engage in and their responses to the characteristics of such tasks. It follows that there are differences in the degree of learning and academic performance that the two mindsets support (Claro et al., 2016; Dweck, 1999).

Theories of Intelligence and Achievement

In studies to date, students with an incremental mindset consistently outperform those espousing an entity view. For example, Romero et al. (2014) surveyed students at four time points from the end of sixth to the end eighth grade. They found that students endorsing incremental beliefs about intelligence had consistently higher grades over the duration of the study. Additionally, students who held these beliefs in sixth grade were significantly more likely to enrol in advanced math courses in seventh and eighth grade, compared to those with fixed beliefs.

Several studies provide evidence suggesting that this disparity in achievement is a behavioral consequence of the different goals, effort attributions, and response patterns between the two mindsets (see Dweck, 1999 for a broad review). For example, in a sample of middle school students, Blackwell, Trzesniewski, and Dweck (2007) found that compared to entity theorists, incremental theorists were more likely to pursue learning goals, believe that high effort led to high performance and think that a change in study strategy could remedy poor performance. Over the two-year duration of their study, these goals and beliefs were accompanied by significant improvements in math grades for incremental theorists, whilst there was no change in grades for entity theorists.

This was supportive of previous research by Hong, Chiu, Dweck, Lin, and Wan (1999), which found that university students with incremental beliefs attributed their task performance significantly more to effort than to ability, whilst those with entity beliefs made significantly more ability attributions. Hong et al. also found that incremental theorists had far greater intentions to pursue remedial action compared to entity theorists (73.3% and 13.3% respectively), when they were told that their poor performance was due to low ability.

Moreover, Sevincer, Kluge, and Oettingen (2014) used a short-term TOI manipulation in a sample of students, to investigate differences in motivational focus between those with entity and incremental beliefs. They found that those in the incremental condition focused significantly more on the future when elaborating on present versus future aspects of their personally relevant academic goals. Sevincer et al. argued that this was a reflection of the ‘growth-mindset’ that is supported by an incremental TOI. They concluded that this future-oriented focus may allow incremental theorists to better recognise and utilise opportunities that support the achievement of desired academic goals.

Given the positive academic outcomes that are associated with incremental beliefs, there has been an increase in interventions promoting ‘growth mindsets’ and ‘cultures of growth’ in educational settings, particularly in the United States of America (Dweck, 2015; Rattan, et al., 2015). These interventions are based on the notion that naïve TOIs can be changed over the medium to long-term. By explicitly teaching students that effort and persistence can improve intelligence by, for example, making new connections in the brain, those who view intelligence as fixed can come to endorse more incremental beliefs (see Dweck, 1999; Blackwell et al., 2007, experiment 2).

However, recent research suggests that there may be a potential downfall to having this mindset. In some instances, people with incremental beliefs have been shown to have poor metacognitive accuracy compared to those with entity beliefs (Miele, et al., 2011; Miele & Molden, 2010). Specifically, they appear to have trouble making accurate JOLs when information is particularly easy, and particularly difficult (Miele et al., 2011). Whilst it is unlikely that sub-optimal monitoring accuracy has the potential to undermine the overall advantages of an incremental

TOI, it remains worthy of investigation, given that metacognitive JOLs inform study behaviour and have implications for academic performance (Metcalfe & Finn, 2008). The following section of this paper provides an explanation of metacognitive judgements, in order to inform the subsequent discussion of recent evidence suggesting that these judgements may operate differently for entity and incremental people.

Metacognitive Accuracy and Delayed JOLs

It is widely accepted that JOLs are a valid index of metacognitive accuracy that represent the perceived likelihood of recall for newly learned information (Rhodes & Tauber, 2011). The cue-utilisation view (Koriat, 1993; 2008) proposes that learners make judgements about the quality of their learning on the basis of a number of mnemonic cues. This includes memory-specific information, such as memory quality, as well as interpretations of processing (encoding) fluency at the time of learning, and the subjective experience of retrieval fluency when information is recalled at a later time (Koriat & Ma'ayan, 2005). JOL accuracy is therefore dependent on the validity of these cues for predicting future recall (Koriat & Ma'ayan, 2005). When they are a valid indicator of future memory, the accuracy of metacognitive judgements will be high, however in the event they are an invalid index of later recallability, the accuracy of such judgements will be reduced (Koriat & Ma'ayan, 2005).

Previous research has indicated that people do rely on the assumption that subjective feelings of ease during learning are indicative of good learning (Koriat 2008). Ease of processing can be due to perceptual factors such as stimulus clarity and exposure time, or conceptual factors such as concept accessibility or coherence of the information (Miele & Molden, 2010). Both perceptual and conceptual fluency make information easier to understand and encode, and have been shown to lead people to judge highly fluent information as better learned than information which is disfluent (Koriat, 2008; Miele & Molden, 2010). Whilst this may be adaptive as a general rule of thumb, high fluency in itself does not guarantee that information will be encoded or recalled; it simply increases the likelihood of this occurring (Finn & Tauber, 2015). However, given that the subjective experience of processing ease (or difficulty) is highly salient immediately after studying new information, it is the primary mnemonic cue that people draw on when making immediate JOLs (Koriat & Ma'ayan, 2005). In the event that processing fluency is not diagnostic of future recall, the often unconscious reliance on this cue when making JOLs can be problematic (Koriat & Ma'ayan, 2005).

In their research investigating influences on metacognitive accuracy, Nelson and Dunlosky (1991) found that delaying JOLs leads to substantial gains in JOL accuracy. In their experiment, participants viewed sets of unrelated English-English word pairs, and gave immediate JOLs for one half of the pairs and delayed JOLs for the other half. The immediate JOLs were given following each individual word-pair presentation, all of which had a viewing time of 8 seconds (for both immediate and delayed). For the delayed JOLs, all pairs were presented first, and then re-presented for JOL elicitation. It was found that delayed JOLs were a significantly better predictor of actual recall performance than immediate JOLs.

Nelson and Dunlosky (1991) suggested that a delay of 30 to 60 seconds (see also Rhodes & Tauber, 2011) between the presentation of information and the elicitation of a JOL, reduces the dependence on processing fluency as a mnemonic cue. They proposed in their *monitoring dual memories* account, that the delay also causes an increased reliance on cues pertaining to the quality of information storage in LTM, such as retrieval fluency (see also Kimball and Metcalfe, 2003). This improves metacognitive accuracy by permitting a more precise assessment of long-term memory (LTM) for the relevant information, thus making delayed JOLs a more accurate predictor of future memory.

Since Nelson and Dunlosky's (1991) pioneering research, the results of numerous studies have provided further evidence in support of the so-called 'delayed-JOL effect'. A meta-analysis by Rhodes and Tauber (2011) found strong support for both the delayed-JOL effect, and the monitoring dual memories account in the current literature. A more focused study by Koriat and Ma'ayan (2005, experiment 1) found the relative contributions of encoding and retrieval fluency to JOLs to be contingent on the timing of JOL elicitation. In this study, immediate JOLs were more strongly correlated with encoding fluency (indexed by self-paced study time) than retrieval fluency (indexed by retrieval latency). Whilst the reverse was true for delayed JOLs, which were more strongly correlated with retrieval fluency. Further, a regression analysis on the data indicated that study time made the strongest contribution to the prediction of immediate JOLs, whilst the contribution of retrieval latency was strongest for delayed JOLs. JOL accuracy also increased with the delay, indicating a delayed JOL effect (Koriat & Ma'ayan, 2005).

Overall, these studies suggest that perceived encoding effort influences immediate JOLs, because it is the most salient cue immediately following the processing of to-be-remembered information. Whereas perceived retrieval effort influences delayed JOLs, as it has the greatest saliency following a recall attempt. Further, not only do the processes involved in generating delayed JOLs correspond more closely with that of a final recall attempt, they are also arguably a more practical measure of real-world study behaviour compared to immediate JOLs. In the majority of instances, people do not study material immediately after processing it, but do so after a delay, whether it be hours, days, weeks or even months. For the purpose of both accuracy and generalisability, it is therefore beneficial to utilise delayed JOLs when investigating study behaviour that occurs outside of controlled laboratory environments.

Binary Vs. Scale JOLs. The most commonly used index to measure JOLs is the percentage probability scale, on which individuals provide values between 0 and 100% that represent the perceived probability of future recall for items of information (Hanczakowski, Zawadzka, Pasek, & Higham, 2013). However, Hanczakowski et al. (2013) provide evidence that this scale measurement may introduce systematic bias, leading to an inaccurate representation of JOL accuracy. Specifically, their results suggest that when assigning JOLs, individuals use the 0-100 scale to rank order items for recallability, but do not make frequency judgements concerning recall. As such, individuals may often fail to interpret this measurement tool in the way that researchers intend.

Hanczakowski et al's. (2013) suggest that a binary JOL response format, in which individuals respond with either yes or no when predicting whether they will recall an item, is a less biased and more realistic measurement. When using binary decisions as an index of metacognitive accuracy, high accuracy is reflected by responding 'yes' to items that are answered correctly on a final recall test, and 'no' to items that are answered incorrectly. This indicates that a person has the ability to discriminate between material that they do, and do not know, and can make accurate predictions of their future memory performance on the basis of this information

(Hanczakowski et al., 2013). Accuracy is reduced if people recall information which they predicted they would not remember, or fail to recall information which they predicted that they would remember. Hanczakowski et al. (2013) termed these errors 'metacognitive misses' and 'metacognitive false alarms' respectively.

Specifically, Hanczakowski et al (2013, experiments 1 and 2) found that a certain metacognitive bias referred to as the underconfidence-with-practice effect (UWP) is evident when JOLs are measured with a probability scale, but not when the JOL format is binary. They further demonstrated (experiment 4) that transforming the probability scale into a binary scale (0-50% equal to a 'no' response and 51-100% equal to a 'yes' response) and re-analyzing the data removed the UWP effect. This result was in line with previous research by Dunlosky et al. (2005), who found that second-order 0-100 (definitely not accurate/definitely accurate) confidence judgements pertaining to first-order scale JOLs did not correspond to JOL values (e.g., participants were not 20% confident that they would recall an item give a JOL of 20%). Instead, higher second-order judgements were given for JOLs at either extreme, suggesting that people may make a yes/no prediction concerning future recall, followed by a determination of their level of confidence in this decision using the percentage scale (Dunlosky et al., 2005).

Hanczakowski et al. (2013) therefore concluded that scale and binary response formats do not always find equivalent effects, even when measuring the same phenomena under the same instructions. Further, they proposed that binary JOLs may be a more valid measure of the early cognitive processes that are used by individuals to self-monitor their learning. As such, effects that extend from scale to binary response formats should be given greater credibility; consistent results across measures suggests that the effects are more likely to represent genuine psychological phenomena (Hanczakowski et al., 2013). Consequently, in the event that an effect does not generalise across measures, it may be appropriate to question its validity (Hanczakowski et al., 2013).

Theories of Intelligence and Metacognitive Accuracy

Recently, research has come to light suggesting that people's JOLs may be influenced by their beliefs about intelligence (Miele et al., 2011; Miele & Molden, 2010). In many learning contexts, people are required to process novel, challenging material, in order to master new concepts and become familiar with new ideas (Miele & Molden, 2010). Resulting decreases in processing fluency that occur from encountering unfamiliar or challenging information necessitate an increase in the amount of effort that one must exert to process and understand information (Miele & Molden, 2010). Given that peoples' beliefs about the meaning of effort are guided, at least in part, by the naïve TOI to which they subscribe (Miele & Molden, 2010), it seems reasonable that their JOLs for new information may also be influenced by these theories.

Previous research, such as that by Koriat (2008), has demonstrated that in a general sense, individuals do use perceptions of processing fluency to inform their JOLs. Koriat found that items studied for longer, and which had more trials to acquisition, were given lower JOLs, and did in-fact have a poorer rate of recall. Koriat concluded that in instances where perceived processing fluency is a valid cue to future memory (i.e., when actual and perceived difficulty correspond), the reliance on this so-called *easily learned, easily remembered (ELER) heuristic* (Koriat, 2008) is beneficial for JOL accuracy. In other words, believing that more effortful processing is associated with a reduced likelihood of future recall represents the reality of knowledge acquisition, provided that the increase in processing effort is due to greater objective difficulty of the information (Koriat, 2008).

However, recent research has provided evidence that entity and incremental beliefs have differential effects on interpretations of processing fluency and subsequent JOL accuracy. Miele et al. (2011) (see also Miele & Molden, 2010) presented a set of 54 Indonesian-English word pairs with differing levels of perceptual similarity (high, moderate, low) to a sample of 75 participants. After viewing each word pair (study time was unlimited), participants gave a JOL on a 0-100 scale that reflected their level of confidence in their ability to recall the English word on a future test, where the Indonesian word would be presented alone. The results of the final recall test indicated that incremental theorists JOLs were highest for difficult, highly engaging items (low similarity) and lowest for easy items (high-similarity). This appeared to reflect a belief that high task engagement signified successful learning, and resulted in under-confident JOLs for easy items and overconfident JOLs for difficult items. This was not the case for entity theorists, whose JOLs were in line with the ELER heuristic, decreasing with corresponding increases in item difficulty, and thus being quite accurate.

The results of this research indicate that for incremental theorists, subjective difficulty in processing may improve the evaluation of a task and its outcome, because the associated increase in effort is perceived as a means of reaching their current learning goals (i.e., is valuable) (see Brinol, Petty & Tormala, 2006; Labroo & Kim, 2009). In contrast, subjective difficulty and effort conflict with entity theorist's performance goals, leading them to negatively evaluate such factors, and positively evaluate ease of processing, which fits with their goals and effort beliefs. Consequently, Miele et al. (2011) suggested that incremental theorists appear to predominantly use a *highly engaged, easily remembered (HEER)* heuristic, and that this can lead to inaccurate JOLs. They concluded that whilst effort and persistence are important, subjective feelings of engagement do not always signify successful learning. In the majority of instances, highly engaging information is so because it is difficult to process and understand, and is therefore less likely to be remembered. Thus, in contrast to the ELER heuristic, the HEER heuristic is unlikely to be (in any instance) a valid indicator of future recallability, posing a problem for the metacognitive accuracy of incremental theorists.

So, whilst the literature overwhelmingly indicates that an incremental mindset is beneficial for academic achievement, Miele et al. (2011) propose that there may be a potential drawback to having incremental beliefs. This is an important proposition that warrants further empirical investigation. Not only to improve our understanding of how this mechanism may operate, but also to determine if there are ways in which metacognitive monitoring can be improved for incremental people.

The Present Study

As such, the present study has three objectives. The first relates to the effect of the HEER heuristic on incremental theorists' JOLs. Theoretical and empirical evidence (Miele et al., 2011) suggests that incremental theorists' interpretations of processing fluency may present a possible problem for their metacognition, and this has implications for study behaviour in real-world settings. However, Miele et al's. (2011) conclusions were based on immediate JOLs. According to the monitoring dual memories account of the delayed-JOL effect (Nelson & Dunlosky, 1991), delayed JOLs are less prone to the influence of processing fluency, and are a more accurate representation of future recall performance. As such, this research aims to examine whether incremental theorist's metacognitive accuracy can be improved by delaying JOLs. Based on the wealth of research demonstrating the advantages of an incremental mindset, we suspect that delaying JOLs may undermine the proposed negative effects of the HEER heuristic on incremental theorist's metacognitive accuracy. This finding would suggest that the heuristic is unlikely to have problematic impacts on metacognition in applied settings. Such a result would be in line with the current literature on the academic achievement of individuals with an incremental TOI.

Second, previous research by Hanczakowski et al., (2013) indicates that the binary JOL format does not always detect the same effects as 0-100 scale judgements. We are therefore interested in determining if Miele et al's. (2011) findings extend to binary JOLs, or whether they may be an artifact of the scale JOL format. To this end, we will employ a response format that uses Yes JOLs to represent a recall prediction, and No JOLs to represent a non-recall prediction.

Third, whilst a literature scan indicates that the delayed-JOL effect has been observed with binary JOLs, it appears that this has not been demonstrated nearly as thoroughly as with scale JOLs. Given that this effect is extremely robust with scale JOLs, and there is at least some evidence to support its existence with a binary response format (Wojcik, Waterman, Lestie, Moulin, & Souchay, 2014), an additional aim of this research is to replicate the delayed-JOL effect with binary JOLs.

Hypotheses. In line with the research objectives, it was hypothesised that the accuracy of incremental theorists' immediate JOLs would be poorer than that of entity theorists, and that this difference would be reduced for delayed JOLs. It was also hypothesised that overall metacognitive accuracy would be better for delayed JOLs than immediate JOLs.

Method

Participants and Design

The sample consisted of 56 participants (46 females, 10 males), aged between 18 and 63 (M = 25.2, SD = 10.05), who were recruited, following ethics approval (Appendix A), from the University of Tasmania and the wider community. Data from an additional 3 participants (original N = 59) was collected, but was omitted from all analyses due to non-compliance with experimental instructions. University students received course credit for their participation, and all other participants were paid 20 dollars. All participants were fluent in English, and confirmed that they had little to no knowledge of Indonesian, or similar languages (i.e., Malay), as Indonesian words were used in the learning task. The study design conformed to a 2 (TOI: entity/incremental) x 2 (timing of JOL: immediate/delayed) x 3 (item difficulty: easy/moderate/difficult) mixed factorial design, with TOI as the between-subjects variable.

Materials and Procedure

As shown in Figure 1, the procedure consisted of five stages. These were completed in the same order by all participants. The presentation of the study material and filler task, and the recording of study choice, study time, and final test accuracy were done using e-prime software. All responses were entered manually by participants, using the computer keyboard.

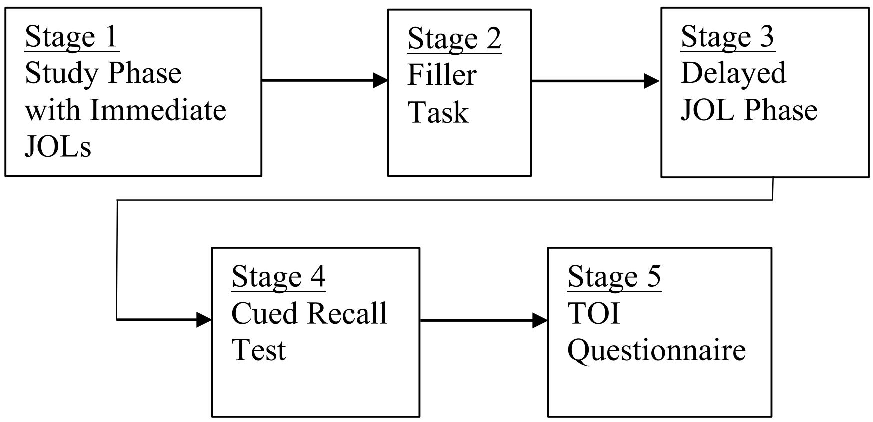

Figure 1. Chronological depiction of the five stages in the procedure.

Study phase with immediate JOLs. Participants gave written informed consent prior to completing any tasks (see Appendix B). They were then seated at a computer and provided with on-screen information (see Appendix C for full transcript), telling them that they would be presented with a series of Indonesian-English word pairs (see Appendix D). Participants were asked to attempt to remember each of the presented word pairs, and told that their memory for them would be tested later in the study. In keeping with Miele et al. (2011, experiment 1), study-time was self-paced, so as to promote high engagement (and thus, the HEER heuristic) by incremental theorists.

Participants were informed that they would only have the opportunity to study each pair once, and that they could not go back and re-study a previous pair after choosing to move on. Order of presentation of the word pairs was random for all participants. Participants were required to press the space bar in order to move to the next screen. Word pairs were presented alone on the screen, and, after choosing to cease studying (i.e., pressing the space bar), participants were asked to make an immediate JOL. This involved responding to a prompt which asked whether they thought they would be able to recall the English word when presented with the Indonesian word alone. Responses were made using the ‘Y’ key, which was used to indicate a recall prediction, and the ‘N’ key which indicated a non-recall prediction.

The word pairs were taken from a paper by Kornell and Son (2009), and were used by Miele et al. (2011). In keeping with Miele et al, 54 of the 96 word pairs were chosen for inclusion. These were split across 3 levels of difficulty (18 pairs in each), to ensure variation in processing fluency. Distinctions between easy, moderate and difficult items were based on correlational norming data provided by Kornell and Son (see Appendix D). For pairs coded as easy, the two words had high perceptual and/or conceptual similarity (e.g., doctor - dokter), for those coded a moderate, most of the words were somewhat similar (e.g., sick - sakit), and for those coded as difficult, the words were largely dissimilar (e.g., live - tinggal).

Filler task. Once immediate JOLs had been provided for all 54 word pairs, participants completed a 2-minute set of basic math problems (e.g., 7 x 5 + 15). They were asked to complete as many as they could in the two minutes. This task was designed to prevent rehearsal of the word pairs and allow this information to move out of STM, so that the following delayed JOLs would be made using predominantly LTM cues.

Delayed JOL phase. After completing the filler task, participants were then presented with the Indonesian word from each pair; presentation order was again random. As for the immediate JOLs, they were required to press the space bar to move to the next screen. Upon doing so, they were again asked to make a Y/N JOL in response to a prompt asking whether they thought they would be able to recall the corresponding English word if presented with the Indonesian counterpart in the future.

Cued recall test. A cued-recall test followed, in which participants were presented with the Indonesian word (in a random order), and asked to recall the matching English word and type their response. The instructions requested that participants make an attempt to recall the target word for each pair, but advised them that if they could not remember a word, they were to leave the space blank and press the space bar to move to the next word. Accuracy of recall responses was scored by two independent raters, and any discrepancies in ratings were discussed and resolved. In instances where a response was clearly misspelled (e.g., resteraunt for restaurant), it was treated as correct, however responses that were conceptually similar but spelled differently (e.g., small for short) were marked as incorrect.

TOI questionnaire. Following the recall test, participants were asked to fill out a pen and paper questionnaire. This comprised a measure of their TOI, in the form of Dweck's (1999) Theories of Intelligence Scale – Self-Form for Adults. This is a well validated, 8-item questionnaire that asks people to rate their agreement from 1 (strongly agree) to 6 (strongly disagree), with each of the statements. Half of the statements are worded to have an incremental focus, such as "you can always substantially change how intelligent you are". Whilst the other half are worded to have an entity focus, such as "you have a certain amount of intelligence, and you can't really do much to change it", with entity items being reverse scored. Scores for all items are summed, with possible scores ranging from 8 (most entity), to 48 (most incremental) (Miele & Molden, 2010). For the purpose of classifying individuals as either entity or incremental, scores can be split at the mid-point (28) of the scale, and those below this value classed as entity, and those above, as incremental (Miele & Molden, 2010). Inspection of scores in the present case indicated that only 10 participants scored below the mid-point. As such, a median split was used in place of a mid-point split, to allow for a more even distribution of participants between the two groups.

The questionnaire also included three language proficiency questions (Appendix E), which were checked prior to data analysis. Although we notified participants that it was a requirement of participation that they have minimal knowledge of Indonesian or Malay language, these are taught in many school curriculums in Tasmania. In addition, the University of Tasmania has a high proportion of International students, and we therefore wanted to be certain that language knowledge would not confound results.

Results

SPSS version 23.0 was used to analyse the main and interaction effects, with alpha levels set at .05 for comparisons of statistical significance. Effect sizes for all pairwise comparisons are reported using Cohen's d, following Cohen's (1988) criteria of 0.2 as a small effect, 0.5 a moderate effect, and 0.8 as a large effect.

Initial inspection of skewness statistics indicated both positive and negative skew within the levels of recall accuracy and metacognitive accuracy, and positive skew within study time. To ensure that this would not undermine the robustness of analyses, all data were transformed (variables exhibiting negative skew were reflected, transformed and then re-reflected) using both square root and natural logarithm transformations, and skewness statistics and histograms re-computed. These indicated that for study time, the log transformations improved the fit to be within an acceptable range. For recall accuracy and metacognitive accuracy, the square root transformation was the superior of the two, however, it did not completely correct the skew, and increased skew for one level of metacognitive accuracy. Analyses were run with both the transformed and untransformed variables, with both providing equivalent results. Thus, for ease of interpretation, all results reported are from the analyses using untransformed data.

In the current study, TOI was treated as a dichotomous variable (Dweck, Chiu, & Hong, 1995; Miele & Molden, 2010, experiment 2) with those below the median score of 33 defined as entity, and those above as incremental (scores ranged from 15 to 46). However, to ensure that any effects found were not an artifact of this approach, data were also analysed using TOI as a continuous moderator variable. Following Miele et al. (2011), this was done by conducting ANCOVA on the DV's, using TOI as a covariate. Using ANCOVA did not change the overall interpretation of any results, and as such, ANOVAs are reported for ease of interpretation.

Effect of Difficulty

Difficulty was not the main variable of interest in the analysis, as the study was based on the established premise that interpretations of processing effort depend on one's naive TOI (Miele et al., 2011). We were mainly interested in whether the effect of this on metacognitive accuracy could be altered by delaying JOLs. However, we still needed to establish whether the difficulty levels ascribed to the sets of word pairs had an effect on both perceptual fluency and memory for the information, to ensure that conclusions made in relation to the hypotheses were validly drawn.

In order to do this, we conducted analyses on both study time, and the accuracy of recall responses. There was a significant effect of difficulty on study time, indicating that participants studied difficult word pairs (M = 5226.45, SD = 3022.69, 95% CI [4421.03, 6031.86]) longer than moderate pairs (M =4695.65, SD = 2336.37, 95% CI [4067.23, 5324.08], d = 0.17) and moderate pairs longer than easy pairs (M = 3072.18, SD = 1247.15, 95% CI [2745.47, 3417.76], d = 0.64), $F(1, 75)$ = 43.32, p < .001 (following a Greenhouse-Geisser correction). There was also a significant effect of difficulty on recall accuracy, with participants recalling more words for the easy pairs (M = 0.78, SD = 0.14, 95% CI [0.74, 0.82]) than moderate pairs (M = 0.29, SD = 0.17, 95% CI [0.25, 0.34], d = 3.04), and more words for moderate pairs than difficult pairs (M = 0.13, SD = 0.17, 95% CI [0.08, 0.18], d = 0.94), $F(1, 86)$ = 590.63, p < .001 (following a Greenhouse-Geisser correction). Together, these results indicate that as expected, the level of both perceived and actual difficulty of the word pairs increased between the easy, moderate, and difficult items.

Effect of Theory of Intelligence x Time Interaction

Metacognitive accuracy. Following Kornell & Rhodes (2013), an index of metacognitive accuracy was obtained by calculating an effect size measure that represented participants' ability to distinguish items they knew from items they did not. As depicted in Figure 2, efficient monitoring is represented by an ability to make recall predictions for items that are recalled at test, and non-recall predictions for items that are not recalled. The effect size used to index efficiency of monitoring was Cohen's w (see Cohen, 1988); a measure designed for use with contingency analyses. It is similar to Cramer's V, with values falling between 0 and 1, and higher values indicating a larger effect (i.e., greater accuracy in present case). The magnitude of the effect is interpreted similarly to that of a correlation, with 0.1 representing a small effect, 0.3 a medium effect, and 0.5 a large effect (Cohen, 1988).

		Correct on Test	
		Yes	No
Predicted to Recall	Yes	X	
	No		X

Figure 2. Accurate metacognitive monitoring is represented by a high degree of match between predicted and actual recall.

Separate effect size values were obtained for each participant at each level of item difficulty, for both immediate and delayed JOLs. However, for a number of participants, there was more than one value of 0 in the contingency table (particularly for the difficult items), meaning that the effect size could not be computed. This considerably reduced the number of data points, and so was corrected following Macmillan and Creelman (1991). This involved adding values of 0.5 to each of the cells containing 0, and for each of these corrections, subtracting 0.5 from one of the cells containing a positive value (i.e., 10, 8, 0, 0 became 9.5, 7.5, 0.5, 0.5).

Initial correlations between participants' JOLs and study time indicated that overall, there was a weak negative correlation between both immediate JOLs and study time, r_{pb}(n = 56) = -0.10, 95% CI [-0.15, -0.05], and delayed JOLs and study time, r_{pb}(n = 56) = -0.14, 95% CI [-0.18, -0.10]. Additional bivariate correlations revealed that neither the immediate JOL-study time correlations or delayed JOL-study time correlations exhibited an association with TOI score (p = .841 and p = .794 respectively). Thus, all participants tended to make lower JOLs as study time (and hence, difficulty) increased. This indicates that both entity and incremental people were relying on the ELER heuristic when making JOLs; a finding that is in contrast to that of Miele et al. (2011).

To test the hypothesis that there would be a significant interaction between the effect of TOI and the effect of time on metacognitive accuracy, a 2 (TOI: entity/incremental) x 2 (timing of JOL: immediate/delayed) x 3 (item difficulty: easy/moderate/difficult) ANOVA was performed on the Cohens *w* values. The predicted interaction was not found, with TOI having no significant effect on either time or difficulty (all alpha values > 0.1)

However, as hypothesised, there was a significant main effect of time on the accuracy of metacognitive judgements, with accuracy being higher for delayed JOLs (M = 0.61 , SD = 0.26, 95% CI [0.56, 0.66]) than immediate JOLs (M = 0.32, SD = 0.20, 95% CI [0.28, 0.36]), F(1, 54) = 124.02, p < .001, d = 1.25, 95% CI [0.89, 1.62], replicating the delayed-JOL effect.

Whilst not central to the main hypothesis of the study, there was also a significant main effect of difficulty on metacognitive accuracy, with accuracy being higher for moderate word pairs (M = 0.53, SD = , 95% CI [0.48, 0.57]) than for easy word pairs (M = 0.47, SD = , 95% CI [0.42, 0.52], d = 0.22), and reducing again for difficult word pairs (*M* = 0.39, *SD* = , 95% CI [0.34, 0.44], *d* = 0.56), $F(2, 108) = 14.44$, $p < .001$.

These main effects were qualified by a significant interaction between time and difficulty, $F(2, 108) = 11.50$, $p < .001$, $\eta p^2 = 0.18$, indicating that the effect of time on metacognitive accuracy was not uniform across the easy, moderate, and difficult word pairs. Table 1 shows the means, SD's, and 95% CI's for the interaction. These values show that when JOLs were immediate, participants were best at accurately determining what they would and would not remember if the word pairs were of a moderate level of difficulty. However, when JOLs were delayed, the accuracy of monitoring for easy and moderate words was on par, with that of the difficult word pairs remaining considerably lower.

Table 1
Mean Accuracy of Immediate and Delayed JOLs at all Three Levels of Item Difficulty.

	Immediate		Delayed	
Difficulty	M (SD)	95% CI	M (SD)	95% CI
Easy	0.27 (0.19)	[0.22, 0.32]	0.67 (0.24)	[0.61, 0.74]
Moderate	0.38 (0.19)	[0.33, 0.43]	0.67 (0.22)	[0.62, 0.74]
Difficult	0.30 (0.21)	[0.24, 0.35]	0.48 (0.26)	[0.41, 0.55]

Paired samples t-tests were conducted to follow-up the effects of time separately for each level of difficulty, all of which were significant following a

Bonferroni adjustment. As illustrated in figure 3, these analyses indicated that the improvement in mean metacognitive accuracy between immediate and delayed JOLs was greatest for the easy word pairs ($M = 0.40$, $SD = 0.26$, 95% CI [0.47, 0.33]), $t(55) = 11.61$, $p < .001$, $d = 1.81$, 95% CI [1.36, 2.25] followed by the moderate word pairs ($M = 0.29$, $SD = 0.26$, 95% CI [0.36, 0.22]), $t(55) = 8.41$, $p < .001$, $d =$ 1.41, 95% CI [1.01, 1.80], being smallest for the difficult word pairs ($M = 0.18$, $SD =$ 0.30, 95% CI [0.26, 0.10]), $t(55) = 4.56$, $p < .001$, $d = 0.78$, 95% CI [0.47, 1.09]. This shows that the delayed-JOL effect was largest for items which were easy to process, and decreased in a linear fashion as items became more difficult.

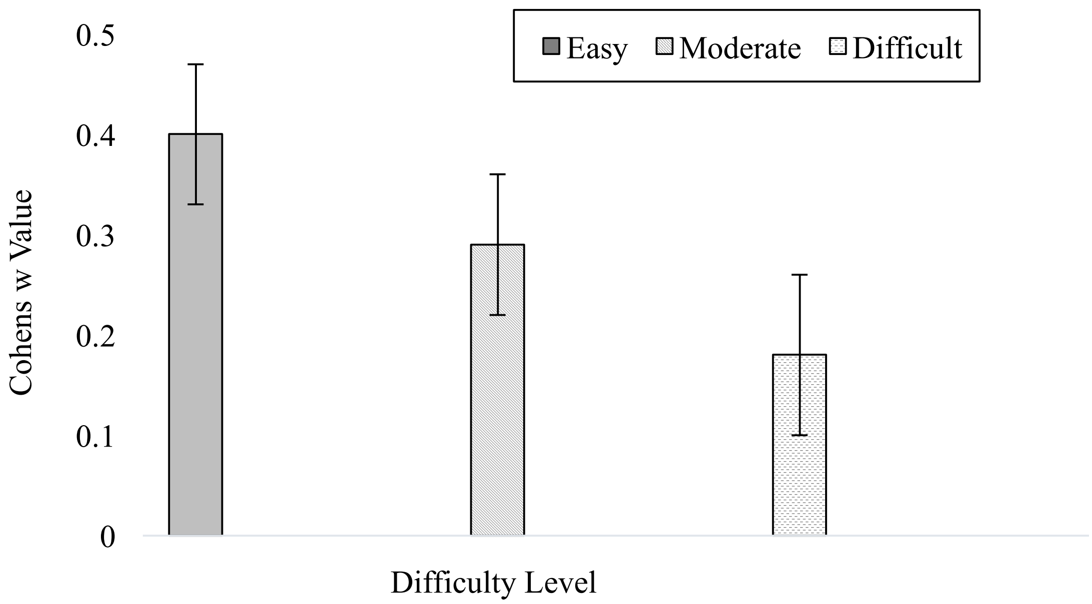

Figure 3. Mean improvement in metacognitive accuracy between immediate and delayed JOLs at each level of word pair difficulty. Error bars show 95% CI's.

Discussion

The primary aim of the present study was to investigate whether delaying JOLs reduced the proposed discrepancy in JOL accuracy between those with an entity versus an incremental TOI (Miele et al., 2011). This research contributes to a limited body of work that has thus far investigated the effects of TOI on JOLs. Our findings do not attest to the idea that entity and incremental theorists rely on different heuristics when making JOLs. Possible reasons for the differences in findings between the current study and that of Miele et al. (2011) are discussed in section 4.4.

Theories of Intelligence and JOL Accuracy

The present study was based on that of Miele et al. (2011), whose findings suggest that individuals with an incremental (but not an entity) TOI interpret processing effort in a way that negatively impacts their metacogntive accuracy. It was hypothesised when JOLs were delayed, and relied predominantly on retrieval effort, the accuracy of incremental theorists' JOLs would be more in line with that of entity theorists. However, this hypothesis was not supported. Contrary to what was expected, there was no difference between the accuracy of entity and incremental theorists' immediate JOLs at any level of item difficulty (i.e., processing effort). Consequently, there was no discrepancy that stood to be improved by delaying JOLs.

Cue validity and heuristics. Whilst these results do not support the notion that incremental theorists utilise a HEER heuristic to guide their metacognitive judgements, they do demonstrate the validity of the ELER heuristic as a basis for JOLs. Our data indicate that study time was positively associated with item difficulty, whilst recall accuracy was negatively associated with difficulty. This is in line with research showing that item difficulty affects the degree of effort needed to process and encode information, and that difficult information is harder to remember (i.e., ELER) (Koriat, 2008).

Further, study time was negatively correlated with JOLs for both entity and incremental theorists, suggesting that individuals use this heuristic to monitor their mastery of information, irrespective of their TOI (Koriat, 2008). It should however be noted that the magnitude of this correlation was weak ($r_{\text{pb}} = -0.1$) and our sample quite incremental (with the median split 5 points above the mid point of the scale). It is therefore possible that incremental theorists use the ELER heuristic to a lesser degree than entity theorists, and may exhibit smaller decreases in JOLs as study time (and thus difficulty) increases. Nonetheless, it is evident that incremental theorists were not making higher JOLs as study time increased (or lower JOLs as it decreased), and were thus not guided by HEER principles.

It should also be noted that the ELER heuristic is only adaptive insofar as the perceived level of task difficulty is diagnostic of actual difficulty (Koriat, 2008), as in the present case. The reliance on perceptual cues, and associated heuristics such as ELER, appears to be a relatively automatic process. As such, previous research has found that manipulating perceived difficulty by altering font size, or sentence coherence (Miele et al., 2011; Miele & Molden, 2010), can negatively impact metacognitive accuracy. Such manipulations do not impact the actual difficulty of the presented information, and can therefore deceive metacognitive judgements (Miele & Molden, 2010). It would be interesting to replicate this study, manipulating only *perceptions* of item difficulty, and observe whether participants show evidence of ELER for immediate judgements, but are able to correct this when JOLs are made at a delay.

In Support of the Delayed-JOL Effect

Whilst it was not the main focus of the present study, the accuracy of participants' JOLs was found to significantly increase when JOLs were made at a delay, supporting the delayed-JOL effect and thus, the second hypothesis. This demonstrates that people are better able to judge what they will and will not remember when these judgements are based on a retrieval attempt, opposed to interpretations of processing fluency. This finding is in line with that of Nelson and Dunlosky (1991), and the extensive body of literature that has since demonstrated the robustness of this effect. Additionally, the present study establishes that the delayed-JOL effect is evident with binary JOLs. This contributes to a seemingly scarce amount of evidence supporting the reliability of this effect with binary judgements (see section 4.3).

Word pair difficulty and monitoring accuracy. The results of the present study did reveal an unusual effect of difficulty on metacognitive accuracy. Research has demonstrated that JOL accuracy tends to be higher for information that is easier to process (see Koriat, Ma'ayan, & Nussinson, 2006, experiment 1). However, in the present study, accuracy was poorer for the easy words than for the moderate words when JOLs were immediate, and no different when JOLs were delayed. This indicates that, even following the improvement in accuracy that resulted from delaying JOLs, there was no advantage in accuracy (i.e., reduced number of metacognitive misses and false alarms) (Hanczakowski et al., 2013) for the word pairs that were easiest to process, when compared to those that were somewhat more difficult. This is contrary to what would be expected based on the literature, and as dicussed in section 4.4, is likely to be the result of a methodological issue with the categorisation of word pairs, not a genuine effect of difficulty.

There was however a standard effect of difficulty for the difficult word pairs, as the increase in accuracy between immediate and delayed JOLs was lowest for these items. This is in line with previous research showing that JOL accuracy reduces as information becomes more effortful to process (Koriat et al., 2006). This indicates that unsurprisingly, learners struggle more when determining how well they will remember novel, challenging information, compared to that which is more closely related to their exisiting knowledge (Metcalfe, 2011).

Binary JOL Format

The results of the present study also lend support to the proposition that the binary JOL format does not always find effects that are evident with a scale format (Hanczakowski et al., 2013). We have no means of direct comparison between binary and scale JOLs in our sample and thus cannot be sure whether a scale JOL measure would have found the same, or different results. However, minimal research has thus far demonstrated evidence of the HEER heuristic, and in instances where it has been detected (Miele et al., 2011; Miele & Molden, 2010), a scale format has been employed. It is of course possible that the scale response format is measuring a true effect that the binary format fails to detect. However, this seems unlikely, given that past research (Dunlosky et al., 2005; Hanczakowski et al., 2013) suggests binary judgements are a more accurate representation of early cognitive processes than scale JOLs.

Further, the delayed-JOL effect has been well established with a scale JOL format, and was clearly evident with the binary format employed here, supporting the idea that robust effects generalise across response formats. It would be beneficial for future research investigating the HEER heuristic to employ both scale and binary JOL formats, to allow comparison across the two measures. If the heuristic is found with scale but not binary JOLs, it would be pertinent to question its validity (Hanczakowski et al., 2013).

Limitations

There are methodological limitations in the current study that may have contributed to the differences in findings between the current study and that of Miele et al. (2011). These are noted as they limit the validity of the current findings, and may provide directions for future research.

Theory of intelligence measure. It was not uncommon for participants to strongly endorse both entity and incremental statements on Dweck's (1999) TOI scale. As such, it is possible that the scale did not provide a true measurement of many participants TOIs. Dweck (1999) acknowledges that the combination of entity and incremental statements can confuse participants, but suggests that this is mostly a concern with school-aged children. A short version of the scale containing only the entity statements (measuring extent of agreement/disagreement) has been validated to address this issue (see Dweck, 1999). As this was not deemed to be a concern in the present study, the complete version of the scale was administered, however, in retrospect, the shortened version may have been more effective. Additonally, this makes it is unclear whether the lack of variability in TOI scores in our sample (i.e., few people endorsing entity views) was an artifact of the TOI measure, or a true lack of variability.

Additionally, whilst both the present study and Miele et al. (2011) used

Dweck's (1999) scale, it it worth noting that Castella and Byrne (2015) propose that a general belief in the malleability of intelligence does not necessarily equate to believing that one's own ability can change. As such, they have recently produced a version of this scale which uses first-person statements (e.g., "I believe I can always substantially improve my intelligence"). Their research demonstrates that this version predicts unique variance in achievement goals and attributions beyond that of Dweck's version, which they propose may measure intelligence beliefs more generally (Castella & Byrne, 2015). Thus, we cannot be certain that participants were responding to the statements in a way that reflected their beliefs about their *own* intelligence.

Categorisation of word pairs. Further, as previously mentioned, there may have been an issue with the categorisation of word pairs that inadvertently improved monitoring accuracy for moderate items. Given that the aim was to replicate Miele et al. (2011), we used the same stimuli with the same three levels of difficulty. However, the moderate category proved quite challenging to classify (Miele et al., 2011 did not provide the words they used for each difficulty level), with few word pairs being of an intermediate level of processing fluency, and the norming data based on a small number of observations. As such, the fluency of some word pairs allocated to the moderate category may have been closer to the 'easy' level, whilst that of others, closer to the 'difficult' level. It is possible that this lead to greater variation between items *within* the moderate category, making discrimination between those that would, and would not be remembered, an easier task, and increasing overall monitoring accuracy for this category. Future research would benefit from using a less ambiguous set of cross-linguistic word pairs, or removing the moderate category and comparing only easy and difficult items.

Goal-driven processing. Finally, in the introductory section of this paper, metacognition was defined as the process of appraising subjective experiences, and using the information gained via this process to evaluate and adjust behaviour (Nelson, 1996). It therefore comprises two main functions; one being the bottom-up monitoring of conscious experience, and the other, the top-down control of adaptive, appropriate responses (Koriat et al., 2006). In the context of the present study, the inherent difficulty (processing fluency) of the presented word pairs acted as a bottom-up influence on monitoring, whilst existing variations in TOI represented a top-down framework, that served to guide the interpretation of this information.

As a primary aim of the present study was to replicate Miele et al. (2011, experiment 1), with the addition of delayed JOLs, we adhered as much as possible to their methodology. This meant manipulating item difficulty to ensure that there were differences in processing fluency to elicit the heuristic mechanim (i.e., HEER) that was of primary interest in this research. However, this also meant that we followed Miele et al's assumption that self-paced study time was sufficient to encourage the active use of implicit goal frameworks (learn versus perform) that form the cornerstone of the incremental and entity TOIs. In other words, we assumed that giving participants unlimited time to do the tasks would ensure that they made genuine attempts at learning the material, and that their interpretations of processing fluency would be guided by their level of task engagement. It is possible that (contrary to Miele et al's assertion), simply allowing study time to be self-paced was not sufficient to encourage true, goal-driven regulation of metacognitive behaviour. If this is so, participants may not have engaged enough in the task to employ the

HEER heuristic.

Future research investigating the HEER heuristic would do well to employ a controlled manipulation, such as a task incentive (i.e., additional payment for higher scores on the final test), to encourage the use of implicit goal frameworks (learn versus perform) (Koriat et al., 2006). Whilst the participants in the current study were allocated research credits or paid for their time, this was not conditional on their task performance and as such, may have not have impacted performance in the way a task-specific incentive would. Whilst Miele et al. (2011) found evidence of the HEER heuristic with self-paced study time and without the use of an incentive, this does not guarantee that the self-paced study method is reliable. It is feasible that if this heuristic does have real practical implications for incremental theorists metacognition, these would be evident in situations which thoroughly engage their cognitive resources, and encourage them to focus on their positive beliefs about effort. We cannot be certain that this occurred in the present case.

Implications

The current research suggests that naïve TOIs do not have differential effects on JOLs. This is not to say that they do not differentially affect other cognitive processes, including values, attiitudes, and motivations. It is clear from the work of many researchers that these factors do change depending on once’s core beliefs about intelligence, and its capacity to grow and change (Blackwell et al., 2008; Claro et al., 2016; Dweck, 1999; Dweck & Leggett, 1988). The present results simply speak to the idea that an incremental TOI does not negatively impact metacognition, and that the positives that stand to be gained from having such a mindset far outweigh any negatives. This has implications for approaches to education, as it suggests that encouraging growth at the policy, classroom, and individual level can have far-reaching positive impacts (see Dweck, 2015).

The present study also demonstrates that learners benefit greatly by delaying their JOLs. This is not a novel finding, but is nonetheless an important one, that supports the utility of study techniques such as distributed practice and self-testing for enhancing learning and memory (Finn & Tauber, 2015). These techniques encourage learners to delay judgements, and promote the use of retrieval practice for informing future study choices (Dunlosky, Rawson, Marsh, Nathan, & Willingham, 2013).

Conclusion

In contrast to Miele et al. (2011), the present study does not support the proposition that individuals with an incremental TOI use the HEER heuristic. The present results instead indicate that the ELER heuristic is used by both entity and incremental theorists, and that it is beneficial for JOL accuracy insofar as it validly indexes item difficulty. The present study also demonstrates that irrespective of TOI, even a slight delay (at least 2 minutes) between the presentation of information and the elicitation of a JOL dramatically improves the accuracy of self-monitoring during learning.

The practical value of these findings lies in informing approaches to learning at the policy, classroom, and individual level. Primarily, this study suggests that an incremental mindset does not damage metacognition, and is thus overwhelmingly beneficial for academic success. As such, it is a worthy use of resources to implement programs and curricula in schools, that focus on individual growth, and foster a belief in the importance of effort in a learning context (Dweck, 2015). This research also speaks to the utility of learning techniques that use retrieval cues as a basis for study behaviour.

References

Blackwell, L.S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development, 78*, 246-263, doi: 10.1111/j.1467-8624.2007.00995.x

Brinol, P., Petty, R. E., & Tormala, Z. L. (2006). The malleable meaning of subjective ease. *Psychological Science, 17*, 200-206. doi: 10.1111/j.1467-9280.2006.01686.x

Castella, K. D., & Byrne, D. (2015). My intelligence may be more malleable than yours: The revised implicit theories of intelligence (self-theory) scale is a better predictor of achievement, motivation, and student disengagement. *European Journal of Psychology of Education, 30*, 245-267. doi: 10.1007/s10212-015-0244-y

Claro, S., Paunesku, D., & Dweck, C. S. (2016). Growth mindset tempers the effects of poverty on academic performance. *Proceedings of the National Academy of Sciences, 113*, 8664-8668. doi: 10.1073/pnas.1608207113

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2^{nd} ed.). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest, 14*, 4-58. doi: 10.1177/1529100612453266

Dunlosky, J., Serra, M. J., Matvey, G., & Rawson, K. A. (2005). Second-Order Judgments About Judgments of Learning. *Journal of General Psychology*, *132*, 335-346. Retrieved from http://web.a.ebscohost.com.ezproxy.utas.edu.au/ehost/pdfviewer/pdfviewer?sid=601b17b4-8cff-4aad-a29c-7d674c174e43%40sessionmgr4010&vid=4&hid=4114

Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality, and development* [electronic-version]. Taylor and Francis; London. Retrieved from http://utas.eblib.com.au.ezproxy.utas.edu.au/patron/FullRecord.aspx?p=1582002

Dweck, C. S. (2015). Growth. *British Journal of Educational Psychology, 85*, 242-245. doi: 10.1111/bjep.12072

Dweck, C. S., Chiu, C., & Hong, Y. (1995). Implicit theories: Elaboration and extension of the model. *Psychological Inquiry*, *6*. 322-333. Retrieved from http://eds.b.ebscohost.com.ezproxy.utas.edu.au/eds/pdfviewer/pdfviewer?sid=326fa0ba-e3f6-493b-8dec-d27a686833fd%40sessionmgr103&vid=1&hid=108

Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review, 95*, 256-273. doi: 10.1037/0033-295X.95.2.256

Finn, B., & Tauber, S. K. (2015). When confidence is not a signal of knowing: How students' experiences and beliefs about processing fluency can lead to miscalibrated confidence. *Educational Psychology Review, 27*, 567-586. doi: 10.1007/s10648-015-9313-7

Hanczakowski, M., Zawadzka, K., Pasek, T., & Higham, P. A. (2013). Calibration of metacognitive judgments: Insights from the underconfidence-with-practice effect. *Journal of Memory and Language, 69*, 429-444. doi: 10.1016/j.jml.2013.05.003

Hong, Y., Chiu, C., Dweck, C. S., Lin, D. M., & Wan, W. (1999). Implicit theories, attributions, and coping: A meaning system approach. *Journal of Personality and Social Psychology, 77*, 588-599. Retrieved from https://psychology.stanford.edu/sites/all/files/Implicit%20Theories,%20Attributions%20and%20Coping_0.pdf

Kimball, D. R., & Metcalfe, J. (2003). Delaying judgements of learning affects memory, not metamemory. *Memory & Cognition, 31*, 918-929. doi: 10.3758/BF03196445

Koriat, A. (1993). How do we know that we know? The accessibility model of the feeling of knowing. *Psychological Review, 100*, 609-639. doi: 10.1037/0033-295X.100.4.609

Koriat, A. (2008). Easy comes, easy goes? The link between learning and remembering and its exploitation in metacognition. *Memory & Cognition, 36*, 416-428. doi: 10.3758/MC.36.2.416

Koriat, A., & Ma'ayan. H. (2005). The effects of encoding fluency and retrieval fluency on judgements of learning. *Journal of Memory and Language, 52*, 478-492. doi: 10.1016/j.jml.2005.01.001

Koriat, A., Ma'ayan, H., & Nussinson, R. (2006). The intricate relationships between monitoring and control in metacognition: Lessons for the cause-and-effect relation between subjective experience and behaviour. *Journal of Experimental Psychology: General, 135*, 36-69. doi: 10.1037/0096-3445.135.1.36

Kornell, N., & Rhodes, M. G. (2013). Feedback reduces the metacognitive benefit of tests. *Journal of Experimental Psychology: Applied, 19*, 1-13. doi: 10.1037/a0032147

Kornell, N., & Son, L. K. (2009). Learners' choices and beliefs about self-testing. *Memory, 17*, 493-501. doi: 10.1080/09658210902832915

Labroo, A. A., & Kim, S. (2009). The "instrumentality" heuristic: Why metacognitive difficulty is desirable during goal pursuit. *Psychological Science, 20*, 127-134. doi: 10.1111/j.1467-9280.2008.02264.x

Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide.* New York: Cambridge University Press.

Metcalfe, J. (2011). Desirable difficulties and studying in the region of proximal learning. In A. S. Benjamin (Ed.), *Successful remembering and successful forgetting: A festschrift in honour of Robert. A. Bjork* (pp. 259-276). Hove, United Kingdom: Psychology Press.

Metcalfe, J., & Finn, B. (2008). Evidence that judgments of learning are causally related to study choice. *Psychonomic Bulletin & Review, 15*, 174-179. doi: 10.3758/PBR.15.1.174

Miele, D. B., Finn, B., & Molden, D.C. (2011). Does easily learned mean easily remembered? It depends on your beliefs about intelligence. *Psychological Science, 22*, 320-324. doi: 10.1177/0956797610397954

Miele, D. B., & Molden, D. C. (2010). Naïve theories of intelligence and the role of processing fluency in perceived comprehension. *Journal of Experimental Psychology: General, 139*, 535-557. doi: 10.1037/a0019745

Nelson, T. O. (1996). Consciousness and metacognition. *American Psychologist, 51,* 102-116. doi: 10.1037/0003-066X.51.2.102

Nelson, T. O., & Dunlosky, J. (1991). When people's judgments of learning (JOLs) are extremely accurate at predicting subsequent recall: The "delayed-JOL effect". *Psychological Science, 2,* 267–270. doi: 10.1111/j.1467-

9280.1991.tb00147.x

Rattan, A., Savani, K., Chugh, D., & Dweck, C. S. (2015). Leveraging mindsets to promote academic achievement: Policy recommendations. *Perspective on Psychological Science, 10*, 721-726. doi: 10.1177/1745691615599383

Rhodes, M. G., & Tauber, S. K. (2011). The influence of delaying judgments of learning on metacognitive accuracy: A meta-analytic review. *Psychological Bulletin, 137,* 131–148. doi:10.1037/a0021705

Romero, C., Master, A., Paunesku, D., Dweck, C. S., & Gross, J. J. (2014). Academic and emotional functioning in middle school: The role of implicit theories. *Emotion, 14*, 227-234. doi: 10.1037/a0035490

Schwarz, N. (2004). Metacognitive experiences in consumer judgment and decision making. *Journal of Consumer Psychology, 14*, 332-348. Retrieved from http://web.a.ebscohost.com.ezproxy.utas.edu.au/ehost/pdfviewer/pdfviewer?vid=5&sid=4057e769-7f30-46f8-a066-fd518ebb7048%40sessionmgr4004&hid=4101

Sevincer, A. T., Kluge, L., & Oettingen, G. (2014). Implicit theories and motivational focus: Desired future versus present reality. *Motivation and Emotion, 38*, 36-46. doi: 10.1007/s11031-013-9359-0

Wojcik, D. Z., Waterman, A. H., Lestie, C., Moulin, C. J. A., & Souchay, C. (2014). Metacognitive judgements-of-learning in adolescents with autism spectrum disorder. *Autism, 18*, 393-408. doi: 10.1177/1362361313479453

Appendices

Appendix A – Ethics approval

UNIVERSITY of TASMANIA
Executive Officer, Social Sciences HREC
Office of Research Services | Research Division
University of Tasmania
Private Bag 1
Hobart TAS 7001
Tel: (03) 6226 2763
[www.utas.edu.au/research]www.utas.edu.au/research]

18 May 2016 3:56 PM

Dr Matthew Palmer
Psychology | School of Medicine, Faculty of Health
Locked Bag 1342

Sent via email

Dear Dr Palmer

Ethics Ref: **H0012660 − Confidence in memory**

This email is to confirm that the following amendment was approved by the Chair of the Tasmania Social Sciences Human Research Ethics Committee on 17/5/2016:

- Addition of researchers Prof Andrew Heathcote, Dr Nicole McCallum, Ms Frances Parkes, and Dr Matthew Gretton.
- Addition of student researchers Valera Griffin, Laura Brumby, Terry Purton, and Caitlin Gleeson.
- Addition of Theories of Intelligence Questionnaire.
- Extension of data collection for a further three years.
- Revised Information Sheet for Participants.

All committees operating under the Human Research Ethics Committee (Tasmania) Network are registered and required to comply with the National Statement on Ethical Conduct in Human Research (NHMRC 2007, updated May 2015).

This email constitutes official approval. If your circumstances require a formal letter of amendment approval, please let us know.

Should you have any queries please do not hesitate to contact me.

Kind regards

Katherine Shaw
Executive Officer, Social Sciences HREC
Office of Research Services | Research Division
University of Tasmania

Appendix B – Participant information sheet and consent form

Judgements of Learning and Memory
Information Sheet for Participants

1. Invitation

We would like to invite you to participate in a psychology experiment about judgements of learning and memory. The experiment is being conducted in partial fulfillment of an honours degree, by University of Tasmania students Caitlin Gleeson and Terry Purton under the supervision of Dr. Matthew Palmer of the Division of Psychology at the University of Tasmania.

2. What is the purpose of this study?

The experiment is investigating factors that affect people's memory for English/foreign language word pairs.

3. Why have I been invited to participate?

For this experiment, we are looking for people aged 18 years or more who have normal or corrected to normal vision (i.e., glasses or contact lenses are fine).

Participation in this study is voluntary – you are entirely free to choose to participate or not, and there will be no consequences if you decide not to participate. If you do participate, any information you provide will be anonymous and no participants in the experiment will be individually identifiable.

4. What will I be asked to do?

Participation would require approximately 1 hour of your time on only one occasion and would take place in a room in the Psychology building on the UTAS campus. The experiment involves viewing a series of word pairs and then answering some questions about them. Participants will also be asked to complete a brief questionnaire about themselves.

5. Are there any possible benefits from participation in this study?

The results of this experiment will help us to understand what factors affect people's memory for a variety of items and events. This information will be useful, for example, in developing better ways to present information in classes. You would be reimbursed for your time with a payment of $20 or 1 hour of research credit.

6. Are there any possible risks from participation in this study?

There are no foreseeable risks associated with participating in this study.

7. What if I change my mind during or after the study?

That's fine - you are free to withdraw from the study at any time, and without providing an explanation. If you choose to withdraw during the study, your responses will be destroyed. If you complete the study, you will not be able to withdraw your data because it will be stored in anonymous form (and so we will not be able to identify which responses are yours).

8. What will happen to the information when this study is over?

The data from this study will be kept in secure storage on the University of Tasmania premises for a period of five years after any publications (e.g., in academic journals) that involve the data. After this period, the data will be archived. Only the researcher will have access to the raw data.
The data will be stored anonymously. All responses will be anonymous and no identifying information will be collected from participants.

9. How will the results of the study be published?

The results of the study will be published in an academic journal. Once the study has been completed, you will be able to access the results by visiting the website below:
http://www.utas.edu.au/psychology/research/research-project-reports
No individual participants will be identifiable in the publication of the results.

10. What if I have questions about this study?

If you have any questions about this study, please feel free to contact us via phone on (03) 6324 3004 (Matthew Palmer) or by email: matthew.palmer@utas.edu.au or cl0@utas.edu.au or tpurton@utas.edu.au

This study has been approved by the Tasmanian Social Sciences Human Research Ethics Committee. If you have concerns or complaints about the conduct of this study, please contact the Executive Officer of the HREC (Tasmania) Network on (03) 6226 7479 or email human.ethics@utas.edu.au. The Executive Officer is the person nominated to receive complaints from research participants. Please quote ethics reference number H0012660.
This information sheet is for you to keep. If you would like to participate in this study, please ask the researcher for a Consent Form to complete.

Thank you for your attention - your time is very much appreciated!

Judgements of Learning and Memory
Participant Consent Form

1. I agree to take part in the research study named above.
2. I have read and understood the Information Sheet for this study.
3. The nature and possible effects of the study have been explained to me.
4. I understand that the study involves viewing a series of word pairs and answering questions about them.
5. I understand that participation involves no foreseeable risks.
6. I understand that all research data will be securely stored on the University of Tasmania premises for five years from the publication of the study results, and will then be destroyed unless I give permission for my data to be archived.

   I agree to have my study data archived. (Note that your data will be stored anonymously.)

   Yes ☐ No ☐
7. Any questions that I have asked have been answered to my satisfaction.
8. I understand that the researchers will maintain confidentiality and that any information I supply to the researcher will be used only for the purposes of the research.
9. I understand that the results of the study will be published so that I cannot be identified as a participant.
10. I understand that my participation is voluntary and that I may withdraw at any time without any effect.

    I understand that I will not be able to withdraw my data after completing the experiment as my data will be anonymous.

Participant's name:
____

Participant's signature:
____

Date: ____________________

Statement by Investigator

☐ I have explained the project and the implications of participation in it to this volunteer and I believe that the consent is informed and that he/she understands the implications of participation.

If the Investigator has not had an opportunity to talk to participants prior to them participating, the following must be ticked.

☐ The participant has received the Information Sheet where my details have been provided so participants have had the opportunity to contact me prior to consenting to participate in this project.

Investigator's name:

____

Investigator's signature:

____

Appendix C – Transcript of participant instructions

Introduction

Welcome to the study. Thank-you for taking the time to participate. Press the SPACE BAR to continue.

During the first part of the study, you will be presented with 54 word pairs. One of the words will be an English word, the other, it's counterpart in Indonesian. Press the SPACE BAR to continue.

Please go through the list in your own time, and do your best to learn each pair. You will only have to opportunity to view each pair once. You will be tested on your memory of each word pair later in the study. Press the SPACE BAR to continue.

Immediate JOL instructions

After studying each word pair, you will be asked whether you think you would be able to recall the English word from the pair, if presented with the Indonesian word only. Press the SPACE BAR to continue.

Please note that once you have made your decision and pressed the space bar, you cannot go back. To begin the study phase, press the SPACE BAR.

Immediate JOL prompt

Do you think you would be able to recall the corresponding English word for this pair?

Y = Yes, N = No. Please use the keyboard to enter your answer, and press the SPACE BAR to continue.

Filler task instructions

The first phase of the study is now complete. During the next part of the study, you will be asked to solve some math problems. Please try to complete as many as you can within the 2 minutes. Press the SPACE BAR when you are ready to begin.

Post-filler task prompt

This phase of the study is now complete. Press the SPACE BAR to continue.

Delayed JOL instructions

During the next part of the study, you will be shown the Indonesian word from each word pair. Please tell us again whether you think you would be able to recall the English word. When you are ready to begin, press the SPACE BAR.

Delayed JOL prompt

Do you think you would be able to recall the corresponding English word for this pair?

Y = Yes, N = No. Please use the keyboard to enter your answer, and press the SPACE BAR to continue.

Cued recall test instructions

During the next part of the study, you will be shown the Indonesian word from each word pair, and asked to RECALL the English word. Use the keyboard to type in your answer, and press the SPACE BAR to move to the next word. Press the SPACE BAR when you are ready to begin.

Recall prompt

Use the keyboard to enter the corresponding English word. Press the SPACE BAR to continue.

Closing statement

The experiment is now finished. Thank-you for your time. If you have any questions or feedback, please see the researcher.

Appendix D – Word pair stimuli

Indonesian-English word pairs taken from Kornell & Son (2009).
The first two columns are the word pairs used. The last two columns contain the norming data, but, as shown, this is based on less than 10 observations.

	Indonesian	English	Accuracy	Observations
Difficult	Terlambat	Late	0	6
	Tinggal	Live	0	6
	Perhiasan	Jewellery	0	6
	Keberangka	Departure	0	7
	Bagaimana	How	0	8
	Sandiwara	Theatre	0	6
	Angin	Wind	0	3
	Pembalut	Bandage	0	7
	Sungai	River	0.12	8
	Sabun	Soap	0.17	6
	Telur	Egg	0.17	6
	Baru	New	0.17	6
	Jelek	Bad	0.2	5
	Basah	Wet	0.2	5
	Duduk	Sit	0.25	8
	Kacamata	Eyeglasses	0.25	8
	Kelapa	Coconut	0.25	8
	Danau	Lake	0.29	7
Moderate	Tinggi	Tall	0.33	6
	Bagasi	Luggage	0.33	6
	Handuk	Towel	0.38	8
	Ombak	Wave	0.38	8
	Debu	Dust	0.38	8
	Rendah	Short	0.4	5
	Panas	Hot	0.43	7
	Restoran	Restaurant	0.43	7
	Sekolah	School	0.43	7
	Jahe	Ginger	0.44	9
	Besar	Big	0.5	6
	Asli	Authentic	0.5	6
	Coro	Cockroach	0.5	8
	Sakit	Sick	0.57	7
	Sutera	Silk	0.57	7
	Reservasi	Reservation	0.6	5
	Pesta	Party	0.6	5
	Sapi	Cow	0.6	5

Easy	Turis	Tourist	0.62	8
	Fotokopi	Photocopy	0.62	8
	Botol	Bottle	0.67	6
	Bon	Bill	0.71	7
	Telepon	Telephone	0.83	6
	Guru	Teacher	0.86	7
	Dokter	Doctor	0.86	7
	Buku	Book	0.86	7
	Gelas	Glass	0.88	8
	Tekstil	Textile	0.88	8
	Foto	Photograph	0.89	9
	Taksi	Taxi	0.89	9
	Wanita	Woman	1	6
	Bis	Bus	1	7
	Polisi	Police	1	8
	Bir	Beer	1	8
	Sama	Same	1	6
	Sandel	Sandal	1	6

Appendix E – Language proficiency questions

Language History Questions

1. What language do you predominantly speak at home?

   ____

2. Can you speak Indonesian fluently? YES / NO (please circle)

3. Have you ever studied Indonesian? YES / NO (please circle)

   3a. For what length of time did you study it (approx.)?

   ____

   3b. How long ago did you study it (approx.)?

   ____

Appendix F

SPSS Data Output

3. difficulty

Measure: MEASURE_1

difficulty	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
1	3081.614	167.665	2745.466	3417.762
2	4695.654	313.448	4067.228	5324.080
3	5226.446	401.728	4421.029	6031.863

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
difficulty	Sphericity Assumed	139047473.497	2	69523736.749	43.322	.000	.445
	Greenhouse-Geisser	139047473.497	1.387	100279757.011	43.322	.000	.445
	Huynh-Feldt	139047473.497	1.438	96706189.851	43.322	.000	.445
	Lower-bound	139047473.497	1.000	139047473.497	43.322	.000	.445
difficulty * TOIgroup	Sphericity Assumed	5020088.260	2	2510044.130	1.564	.214	.028
	Greenhouse-Geisser	5020088.260	1.387	3620441.409	1.564	.219	.028
	Huynh-Feldt	5020088.260	1.438	3491423.441	1.564	.218	.028
	Lower-bound	5020088.260	1.000	5020088.260	1.564	.216	.028
Error(difficulty)	Sphericity Assumed	173320611.635	108	1604820.478			
	Greenhouse-Geisser	173320611.635	74.876	2314763.491			
	Huynh-Feldt	173320611.635	77.643	2232274.632			
	Lower-bound	173320611.635	54.000	3209640.956			

a. Computed using alpha = .05

3. difficulty

Measure: MEASURE_1

difficulty	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
1	.780	.019	.742	.819
2	.292	.023	.245	.339
3	.131	.023	.084	.178

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
difficulty	Sphericity Assumed	12.750	2	6.375	590.625	.000	.916
	Greenhouse-Geisser	12.750	1.591	8.015	590.625	.000	.916
	Huynh-Feldt	12.750	1.662	7.673	590.625	.000	.916
	Lower-bound	12.750	1.000	12.750	590.625	.000	.916
difficulty * TOIgroup	Sphericity Assumed	.001	2	.000	.037	.964	.001
	Greenhouse-Geisser	.001	1.591	.001	.037	.936	.001
	Huynh-Feldt	.001	1.662	.000	.037	.942	.001
	Lower-bound	.001	1.000	.001	.037	.848	.001
Error(difficulty)	Sphericity Assumed	1.166	108	.011			
	Greenhouse-Geisser	1.166	85.904	.014			
	Huynh-Feldt	1.166	89.729	.013			
	Lower-bound	1.166	54.000	.022			

a. Computed using alpha = .05

Descriptive Statistics

	Mean	Std. Deviation	N
IJOL Point biserial correlation	-.100996358998295	.198363696975509	56
TOI Score	32.73	6.496	56

Correlations

		IJOL Point biserial correlation	TOI Score
IJOL Point biserial correlation	Pearson Correlation	1	.027
	Sig. (2-tailed)		.841
	N	56	56
TOI Score	Pearson Correlation	.027	1
	Sig. (2-tailed)	.841	
	N	56	56

Descriptive Statistics

	Mean	Std. Deviation	N
TOI Score	32.73	6.496	56
DJOL Point biserial correlation	-.140209901483276	.164396690385745	56

Correlations

		TOI Score	DJOL Point biserial correlation
TOI Score	Pearson Correlation	1	-.036
	Sig. (2-tailed)		.794
	N	56	56
DJOL Point biserial correlation	Pearson Correlation	-.036	1
	Sig. (2-tailed)	.794	
	N	56	56

3. difficulty

Measure: MEASURE_1

difficulty	Mean	Std. Error	95% Confidence Interval Lower Bound	Upper Bound
1	.470	.024	.422	.518
2	.528	.022	.484	.572
3	.391	.024	.342	.440

5. time

Measure: MEASURE_1

time	Mean	Std. Error	95% Confidence Interval Lower Bound	Upper Bound
1	.317	.019	.278	.355
2	.610	.025	.559	.661

7. time * difficulty

Measure: MEASURE_1

time	difficulty	Mean	Std. Error	95% Confidence Interval Lower Bound	Upper Bound
1	1	.271	.026	.218	.323
	2	.381	.026	.329	.433
	3	.298	.028	.242	.354
2	1	.670	.032	.605	.735
	2	.676	.030	.615	.736
	3	.484	.035	.413	.554

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
time	Sphericity Assumed	7.182	1	7.182	124.017	.000	.697
	Greenhouse-Geisser	7.182	1.000	7.182	124.017	.000	.697
	Huynh-Feldt	7.182	1.000	7.182	124.017	.000	.697
	Lower-bound	7.182	1.000	7.182	124.017	.000	.697
time * TOIgroup	Sphericity Assumed	.053	1	.053	.920	.342	.017
	Greenhouse-Geisser	.053	1.000	.053	.920	.342	.017
	Huynh-Feldt	.053	1.000	.053	.920	.342	.017
	Lower-bound	.053	1.000	.053	.920	.342	.017
Error(time)	Sphericity Assumed	3.127	54	.058			
	Greenhouse-Geisser	3.127	54.000	.058			
	Huynh-Feldt	3.127	54.000	.058			
	Lower-bound	3.127	54.000	.058			
difficulty	Sphericity Assumed	1.063	2	.531	14.439	.000	.211
	Greenhouse-Geisser	1.063	1.895	.561	14.439	.000	.211
	Huynh-Feldt	1.063	1.999	.532	14.439	.000	.211
	Lower-bound	1.063	1.000	1.063	14.439	.000	.211
difficulty * TOIgroup	Sphericity Assumed	.000	2	.000	.003	.997	.000
	Greenhouse-Geisser	.000	1.895	.000	.003	.996	.000
	Huynh-Feldt	.000	1.999	.000	.003	.997	.000
	Lower-bound	.000	1.000	.000	.003	.954	.000
Error(difficulty)	Sphericity Assumed	3.974	108	.037			
	Greenhouse-Geisser	3.974	102.347	.039			
	Huynh-Feldt	3.974	107.926	.037			
	Lower-bound	3.974	54.000	.074			
time * difficulty	Sphericity Assumed	.638	2	.319	11.495	.000	.176
	Greenhouse-Geisser	.638	1.957	.326	11.495	.000	.176
	Huynh-Feldt	.638	2.000	.319	11.495	.000	.176
	Lower-bound	.638	1.000	.638	11.495	.001	.176
time * difficulty * TOIgroup	Sphericity Assumed	.002	2	.001	.038	.962	.001
	Greenhouse-Geisser	.002	1.957	.001	.038	.960	.001
	Huynh-Feldt	.002	2.000	.001	.038	.962	.001
	Lower-bound	.002	1.000	.002	.038	.845	.001
Error(time*difficulty)	Sphericity Assumed	2.997	108	.028			
	Greenhouse-Geisser	2.997	105.660	.028			
	Huynh-Feldt	2.997	108.000	.028			
	Lower-bound	2.997	54.000	.055			

a. Computed using alpha = .05

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Cohens w_DJOL_easy	.668670161685919	56	.240952121436570	.032198581608286
	Cohens w_IJOL_easy	.271519389811866	56	.193585964651452	.025869012664781
Pair 2	Cohens w_DJOL_moderate	.674746471202873	56	.223645965401768	.029885949230990
	Cohens w_IJOL_moderate	.381831641546127	56	.191107599979551	.025537827254719
Pair 3	Cohens w_DJOL_difficult	.482805987918358	56	.260009694988316	.034745256995924
	Cohens w_IJOL_Difficult	.298827127565724	56	.208241115141132	.027827389524924

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper			
Pair 1	Cohens w_DJOL_easy - Cohens w_IJOL_easy	.397150771874053	.255942595857124	.034201768013532	.328608897107280	.465692646640825	11.612	55	.000
Pair 2	Cohens w_DJOL_moderate - Cohens w_IJOL_moderate	.292914829656746	.260792501340499	.034849863894855	.223074141719938	.362755517593554	8.405	55	.000
Pair 3	Cohens w_DJOL_difficult - Cohens w_IJOL_Difficult	.183978860352635	.301957514038135	.040350770103310	.103114110025411	.264843610679858	4.559	55	.000